

Henryk PRZYBYŁA

WYBRANE PROBLEMY WIELOWYMIAROWEJ ANALIZY PORÓWNAWCZEJ  
W ODNIESIENIU DO ZAGADNIEŃ GÓRNICZYCH

**Streszczenie.** W artykule zaprezentowano podstawowe problemy wielowymiarowej analizy porównawczej zwracając szczególną uwagę na przygotowanie zbioru zmiennych opisujących badany obiekt w aspekcie celu analizy, sposób normowania zmiennych, procedurę wyboru podziału. Rozważania teoretyczne uzupełnione zostały przykładem zaczerpniętym z praktyki górniczej.

## 1. WSTĘP

W ostatnich latach obserwuje się wzrost zainteresowania wielowymiarową analizą porównawczą, przy czym zainteresowanie to dotyczy zarówno rozwoju metod stosowanych w tego typu analizie jak również dziedzin jej zastosowania. W przekonaniu autora górnictwo węglowe jest szczególnie predysponowane do korzystania z wielowymiarowej analizy porównawczej, szczególnie ze względu na trzy jego cechy, tj.:

- znaczący wpływ warunków naturalnych na wyniki produkcyjne i ekonomiczne,
- losowy charakter wyników produkcyjnych i ekonomicznych,
- konieczność wielokryterialnej oceny wyników pracy kopalni.

Celem pobudzenia zainteresowania wielowymiarową analizą porównawczą w odniesieniu do zagadnień górniczych w niniejszym opracowaniu przedstawiono podstawy teoretyczne samej analizy poparte przykładem zaczerpniętym z praktyki górniczej.

## 2. PODSTAWY TEORETYCZNE WIELOWYMIAROWEJ ANALIZY PORÓWNAWCZEJ

Przedmiotem wielowymiarowej analizy porównawczej mogą być przedsiębiorstwa górnicze (kopalnie węgla kamiennego, Przedsiębiorstwa Robót Górniczych, Zakłady Remontowo-Budowlane itp.) lub ich części, lub to samo przedsiębiorstwo w różnych okresach czasu (np. latach). Wielowymiarowa analiza porównawcza może obejmować całokształt działalności przedsiębiorstw lub wybrane fragmenty ich działalności, np. zatrudnienia, robót przygotowawczych itd.

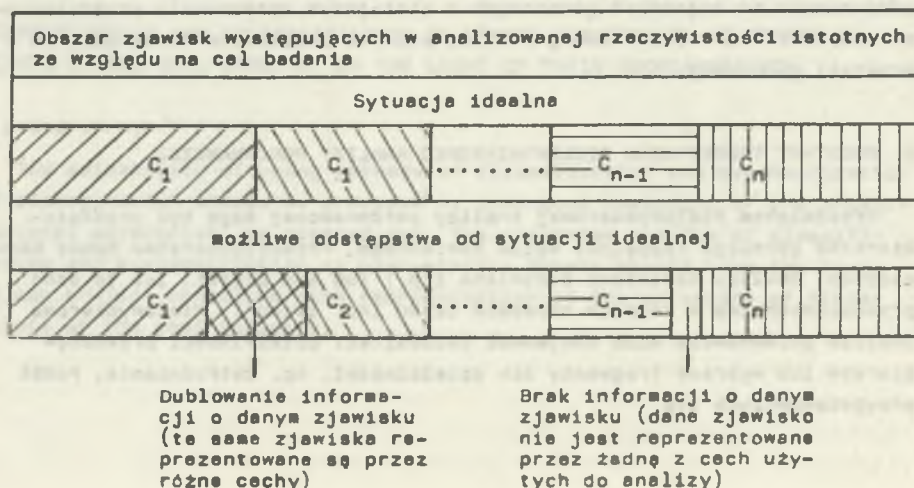
Przy stosowaniu wielowymiarowej analizy porównawczej wyróżnić można trzy etapy prac, tj.:

1. Prace wstępne, w ramach których formułuje się cel i zakres analizy, określa się stopień szczegółowości informacji i sposoby jej pozyskiwania, dokonuje się wstępnej obróbki danych statystycznych.
2. Prace obliczeniowe, w ramach których dokonuje się wyboru algorytmów obliczeniowych, dobiera się programy obliczeniowe i wykonuje obliczenia.
3. Prace końcowe, w ramach których formułuje się wnioski końcowe i przeprowadza się ich analizę i interpretację.

## 2.1. Zakres prac wstępnych

Dla sformułowanego celu analizy dobieramy zakres informacji jaki jest niezbędny do opisu badanej rzeczywistości w aspekcie celu analizy. Jest to zdaniem autora najważniejszy a zarazem i najtrudniejszy moment w całej analizie. Stosowane algorytmy umożliwiają otrzymanie wyniku co nejwyżej na tyle wiarygodnego, na ile wiarygodnie wprowadzone informacje opisują badaną rzeczywistość. Wiadomo, że dla pełnego opisu rzeczywistości należy użyć nieskończonej liczby cech; w badaniach stosowanych ograniczamy się tylko do cech istotnych ze względu na cel analizy i to tylko do tych cech, które potrafimy wyrazić za pomocą liczb. Redukcję liczby cech umożliwiał nam również agregacja informacji. Agregacja bądź dezagregacja informacji uzależniona jest jednak od wymagań co do szczegółowości uzyskiwanych wyników.

W większości publikacji przy doborze cech podkreśla się znaczenie analizy formalnej opartej na metodach statystycznych pomijając niemal całkowicie analizę merytoryczną opartą na zakresach zjawisk jakie te cechy reprezentują. Wyjątek stanowią prace J. Więckowskiego, np. [11]. Ideę analizy merytorycznej można przedstawić w sposób graficzny jak poniżej:



Ideę redukcji liczby cech przedstawić można graficznie w sposób następujący:

## egregacja

|  |              |                |                  |
|--|--------------|----------------|------------------|
|  | powierzchnia | produkcyjne    | środki<br>trwałe |
|  | dół          |                |                  |
|  | powierzchnia | nieprodukcyjne |                  |
|  | dół          |                |                  |

— dezagregacja

Prezentowany podział środków trwałych jest tylko jednym z możliwych podziałów, podobnie ma się z innymi cechami, stąd przygotowany materiał statystyczny (zbiory cech) poddajemy analizie formalnej - statystycznej. Wśród różnych metod polecić można metodę Z. Hellwiga, metody graficzne, metodę analizy czynnikowej, analizę współczynników zmienności i mierników wartości informacyjnej cech.

Przy korzystaniu z analizy współczynnika zmienności i wartości informacyjnej cech korzystamy z poniższych zależności:

$V_j$  - współczynnik zmienności,

$W_j$  - miernik względnej wartości informacyjnej cech.

$$V_j = \frac{s_j}{\bar{c}_j}$$

$$s_j = \frac{1}{n} \left[ \frac{1}{m} \sum_{i=1}^m (c_{ij} - \bar{c}_j)^2 \right]^{1/2}$$

$$\bar{c}_j = \frac{1}{m} \sum_{i=1}^m c_{ij}$$

$$W_j = \frac{V_1}{\sum_{j=1}^n V_j}$$

Ze zbioru zmiennych wybieramy te, które charakteryzują wysokie współczynniki zmienności i wartości informacyjnej cech. Przygotowany materiał grupujemy w dwóch macierzach: macierzy nakładów i macierzy wyników.

W macierzy nakładów zawarte są cechy informujące o warunkach górniczo-geologicznych, stosowanej technologii, technice i organizacji pracy. Macierz wyników zawiera informacje o wynikach produkcyjnych i ekonomicznych.

Macierz nakładów wykorzystamy do grupowania obiektów (obiektami mogą być Przedsiębiorstwa Robót Górniczych, kopalnie lub ich części itd.) w podzbiory statystycznie jednorodne.

Grupowanie obiektów w podzbiory statystycznie jednorodne czyni zadość wymaganiom analizy porównawczej - porównywać można jedynie obiekty porównywalne.

Ponieważ każda z cech w macierzach nakładów i wyników wyrażona jest w odpowiedniej dla siebie jednostce, zachodzi konieczność ich normowania. Prawidłowo przeprowadzone normowanie powinno spełniać następujące postulaty:

- wartości unormowane są liczbami niemianowanymi, niezależnie od rodzaju cech, których wartości są unormowane,
- wartości unormowane są nieujemne,
- wartości należą do skończonego, uniwersalnie unormowanego przedziału liczbowego,
- w miarę wiernie odzwierciedlają zmienność wartości nieunormowanej.

Normowanie możemy przeprowadzić przez:

- przekształcenie ilorazowe

$$x_1 = \frac{c_1 \min}{c_1} \quad \text{lub} \quad x_1 = \frac{c_1}{c_1 \max},$$

- standaryzację

$$x_1 = \frac{c_1 - \bar{c}_1}{s_{c_1}} \quad \text{gdzie} \quad \bar{c}_1 = \frac{1}{m} \sum_{i=1}^m c_1$$

$$s_{c_1} = \left[ \frac{1}{m} \sum_{i=1}^m (c_1 - \bar{c})^2 \right]^{1/2},$$

- normalizację

$$x_1 = \frac{c_1}{s_{c_1}}.$$

Wśród tych metod wybieramy tę, która spełnia warunek:

$$V_{z1} = \frac{S_{z1}}{z1} \rightarrow \min,$$

gdzie:

$$z_1 = \frac{c_1}{x_1}; \quad \bar{z}_1 = \frac{1}{n} \sum_{i=1}^n z_1;$$

$$s_{z_1} = \left[ \frac{1}{n} \sum_{i=1}^n (z_1 - \bar{z}_1)^2 \right]^{1/2}.$$

## 2.2. Prace obliczeniowe

Wśród prac obliczeniowych występują: podział zbioru obiektów na podzbiory statystycznie jednorodne, wyznaczenie syntetycznych miar rozwoju obiektów (obliczenia te możemy ograniczyć do obiektów w ramach podzbioru, w którym występuje badany obiekt), wyznaczenie stopnia podobieństwa obiektów i wyznaczenie stopnia podobieństwa cech, budowa modeli ekonometrycznych istotnie różnych cech (obliczenia te prowadzimy tylko dla cech zawartych w macierzy wyników).

Przed przystąpieniem do obliczeń musimy określić liczbę obiektów, które będą uczestniczyły w analizie oraz czasokres obserwacji. Liczba ta uzależniona jest od liczby obiektów istniejących oraz od zakładanego poziomu istotności wyników.

Uważam, że również nie bez znaczenia jest czas- i pracochłonność pozyskiwania informacji oraz koszt obliczeń numerycznych. Zwiększenie liczby obiektów możemy uzyskać zbierając informacje o każdym z obiektów w różnych momentach czasowych. Wybierając czasokres obserwacji (doba, miesiąc, kwartał) należy przeanalizować zmienność cech i wybrać taki czasokres, dla którego współczynniki zmienności są możliwie najmniejsze.

Na przykład analizując wielkość wydobycia ze ściany

|      |            |            |
|------|------------|------------|
| dla: | t = 1 doba | V = 18,06% |
|      | t = 2 doby | V = 14,49% |
|      | t = 3 doby | V = 14,07% |
|      | t = 4 doby | V = 12,11% |
|      | t = 5 dob  | V = 10,84% |
|      | t = 6 dob  | V = 10,41% |

do analizy wybieramy  $t = n \cdot 6 \quad n = 1, 2, 3, \dots$

### 2.2.1. Podział zbioru obiektów na podzbiory statystycznie jednorodne

Stosując do podziału zbioru obiektów na podzbiory statystycznie jednorodne wg jednego z algorytmów prezentowanych w pracach [3, 4, 10] otrzymujemy szereg podziałów i zachodzi konieczność wyboru jednego z nich. Ze wszystkich możliwych podziałów wybieramy ten, dla którego zachodzi nierówność:

$$F > F_{\alpha} \quad \frac{F}{F_{\alpha}} \rightarrow \max,$$

gdzie:

$$F = \frac{s_{kg}^2}{s_{wg}^2}$$

$s_{kg}^2$  - wariancja międzygrupowa,

$s_{wg}^2$  - wariancja wewnątrzgrupowa.

Przy wyborze podziału można również kierować się miarą wewnętrzną spójności obiektów  $w^{zt}$

$$w^{zt} = \frac{2}{s(s-1)} \sum d_{kk'},$$

gdzie:

$$d_{kk'} = \frac{1}{m} \sum_{j=1}^m |y'_{kj} - y'_{k'j}|$$

$s$  - liczba obiektów  $P_k$  w grupie  $zt$ ,

$y'_{kj}$  - unormowane wartości cech w macierzy wyników.

## 2.2.2. Wyznaczenie syntetycznej miary rozwoju

W pierwszej kolejności wyznaczamy abstrakcyjny punkt  $P_0$  o współrzędnych  $y_{01}, y_{02}, \dots, y_{0m}$  spełniających warunki:

$$y_{0j} = \max_i y_{ij} \quad \text{gdy } j \in s \quad (s - \text{zbiór cech stymulant})$$

$$y_{0j} = \min_i y_{ij} \quad \text{gdy } j \in D \quad (D - \text{zbiór cech destymulant})$$

Odległości pomiędzy poszczególnymi punktami  $P_i$  ( $i = 1, 2, \dots, N$ ) a punktem  $P_0$  obrazującym wzorzec rozwoju oblicza się według wzoru:

$$c_i = \left[ \sum_{j=1}^m (y_{ij} - y_{0j})^2 \right]^{1/2}$$

Syntetyczną miarę rozwoju  $d_i$  oblicza się z poniższych zależności:

$$d_i = 1 - \frac{c_i}{c_0} \quad c_0 = \bar{c}_0 + 2s_0$$

$$s_0 = \left[ \frac{1}{N} \sum_{i=1}^N (c_{i0} - c_0)^2 \right]^{1/2}$$

$$c_0 = \frac{1}{N} \sum_{i=1}^N c_{i0}$$

Obliczenia syntetycznej miary rozwoju możemy ograniczyć do podzbioru, w którym występuje obiekt badany.

### 2.2.3. Wyznaczanie stopnia podobieństwa obiektów i cech

Dla dwóch obiektów  $O_k$ ;  $O_k^2$   $k_1, k_2 = 1, 2, \dots, k$ ;  
 $k_1 \neq k_2$  miarę podobieństwa wyznaczamy z zależności:

$$dk_1 k_2 = \frac{1}{m} \sum_{j=1}^m |y_{k_1 j}^1 - y_{k_2 j}^2|$$

gdzie:

$m$  - liczba cech unormowanych w macierzy wyników.

U w a g a: Dla cech w macierzy wyników normowanie prowadzimy według przekształceń ilorazowych:

$$y_{kj} = \frac{c_{kj}}{\max c_{kj}} - \text{dla cech stymulant,}$$

$$y_{kj} = \frac{\min c_{kj}}{c_{kj}} - \text{dla cech destymulant.}$$

Dla dwóch wyróżnionych wektorów cech unormowanych  $x_p$  i  $x_g$  miarę podobieństwa określamy z zależności:

$$r_{pg} = \frac{1}{N_1} \sum_{i=1}^{N_1} x_{ip} \cdot x_{ig}$$

U w a g a. Dla wyznaczania cech podobnych i cech istotnie różniących się (przy zadanym poziomie istotności  $\alpha$ ) możemy korzystać z analizy wariancji.

Przykład

Kopalnia węgla kamiennego "B" jest zainteresowana analizą stanu zatrudnienia. Zbiór cech istotnych ze względu na stan zatrudnienia macierzy nakładów stanowią:

- wielkość wydobycia,  $T/m-c$
- średni postęp ścian,  $m/m-c$
- średnia długość frontu górniczego,  $m$
- liczba pokładów w eksploatacji,
- liczba oddziałów wydobywczych,
- węgiel netto wybrany mechanicznie,  $T/m-c$
- liczba pracujących kombajnów ścianowych,
- liczba pracujących kombajnów chodnikowych,
- długość ciągów taśmowych,  $m$
- długość wyrobisk chodnikowych,  $m$ .

Przy ustalaniu listy cech analizę merytoryczną uzupełniono analizą mierników wartości informacyjnej cech.

W macierzy wyników zbiór cech stanowiły:

- wydajność pracy,  $T/pdn$
- koszty wydobycia,  $zł/T$
- średnia cena zbytu,  $zł/T$
- wskaźnik wypadkowości,
- współczynnik czystości wybierania,
- wskaźnik wykonania planu wydobycia.

Przy normowaniu cech w macierzy nakładów korzystano ze standaryzacji, natomiast normowanie cech w macierzy wyników oparto na przekształceniach ilorazowych.

Do podziału zbioru kopalń na podzbiory statystycznie jednorodne korzystano z algorytmu prezentowanego w pracy [4] a do analizy wybrano podział, przy którym stosunek  $F$  do  $F_{0,05}$  był maksymalny.

W wyniku podziału otrzymano pięcioelementowy podzbiór kopalń, w którym występowała kopalnia "B". W prezentowanej tablicy 1 przedstawiono interesujące kopalnię stany zatrudnienia. Z przeprowadzonych obliczeń wynika, że kopalnia "B" powinna mieć stany zatrudnienia zbliżone do kopalni "C". Dodajmy, że kopalnia "C" o najwyższym wskaźniku podobieństwa ma również wyższy miernik rozwoju niż kopalnia "B", który uzyskała przy znacząco mniejszym stanie zatrudnienia. Wśród cech istotnych ze względu na stan zatrudnienia znacząco różniły się

|                                     | KWK "B" | KWK "C" |
|-------------------------------------|---------|---------|
| postęp ścian                        | 563     | 702     |
| węgiel brutto wybrany maszynowo     | 245383  | 329183  |
| liczba kombajnów chodnikowych       | 5       | 2       |
| przenośniki taśmowe, długość ciągów | 10718   | 21296   |

Tablice 1

Zbiórce zestawienie cech opisujących stan zatrudnienia w kopalniach należących do tego samego podzbioru w przeliczeniu na 1000 t wydobywania

| Lp. | Kopalnie węgle kamiennego | Mierniki rozwoju x1 | Wskaźniki podobieństwa KWK "A" | Przeciętna liczba zatrudnionych w wyrobiskach wybiernych | Przeciętna liczba zatrudnionych robotników grupy przemysłowej i rozwojowej | Przeciętna liczba zatrudnionych - grupa nieprzemysłowa | Przeciętna liczba zatrudnionych - ogółem kopalnie wydobywcze | Stan zatrudnienia |
|-----|---------------------------|---------------------|--------------------------------|----------------------------------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------|--------------------------------------------------------------|-------------------|
| 1   | 2                         | 3                   | 4                              | 5                                                        | 6                                                                          | 7                                                      | 8                                                            | 9                 |
| 1   | KWK "A"                   | 0,438               | 0,87                           | 178                                                      | 517                                                                        | 38                                                     | 647                                                          | 665               |
| 2   | KWK "B"                   | 0,402               | -                              | 188                                                      | 516                                                                        | 70                                                     | 680                                                          | 700               |
| 3   | KWK "C"                   | 0,417               | 0,91                           | 141                                                      | 505                                                                        | 51                                                     | 640                                                          | 658               |
| 4   | KWK "D"                   | 0,388               | 0,74                           | 202                                                      | 557                                                                        | 56                                                     | 717                                                          | 732               |
| 5   | KWK "E"                   | 0,416               | 0,82                           | 153                                                      | 444                                                                        | 25                                                     | 593                                                          | 607               |

x1 - Mierniki rozwoju były liczone dla całego zbioru kopalń. Inne wielkości otrzymano by, gdyby mierniki te liczone były tylko dla wyróżnionego podzbioru.

Można zatem powiedzieć, że kopalnia "C" uzyskała lepsze wyniki produkcyjne i ekonomiczne dzięki większej mechanizacji robót korytarzowych oraz lepszej organizacji pracy w wyrobiskach wybierkowych.

#### ZAKOŃCZENIE

Zaprezentowana w sposób skrótowy wielowymiarowa analiza porównawcza może być wykorzystywana w ocenie pracy całej kopalni jak również w wybranych dziedzinach jej działalności. Wyniki analizy służą ocenie stanu istniejącego przez porównanie własnych wyników z wynikami pracy kopalń wykazujących najwyższe podobieństwo z badaną kopalnią. Uważam, że stosowanie wielowymiarowej analizy porównawczej poprzez jej związek z psychologiczną teorią decyzji, w której podkreśla się znaczenie poziomu aspiracji decydenta jako kryterium wyboru takiej, a nie innej decyzji, przyczynić się będzie do wzrostu poziomu aspiracji a tym samym do racjonalizacji decyzji. Przy interpretacji wyników obliczeń może być stosowane podejście nadlerowskie przy którym analizuje się nie przyczyny odchyień, lecz działania, jakie należy podjąć, aby zbliżyć się do rozwiązania wzorcowego. Rozwiązanie wzorcowe stanowi bowiem rozwiązanie z punktu widzenia technicznego, technologicznego i organizacyjnego możliwe do realizacji.

#### LITERATURA

- [1] Bartosiewicz S.: Propozycje metody tworzenia zmiennych syntetycznych. Prace A.E. Wrocław 84, 1986.
- [2] Borys T.: Metody normowania cech. P.S. 2, 1978.
- [3] Bureki J., Ziembicki L.: O pewnej klasyfikacji zbiorów. P.S. 2, 1974.
- [4] Czabanka J.: Symulacyjny model zarządzania kopalnią węgla kamiennego. Praca doktorska Pol. Śl., Gliwice 1979.
- [5] Hellwig Z.: Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju. P.S. 4, 1968.
- [6] Kozyra J.: Typologiczny podział kopalń ze względu na wskaźniki techniczno-organizacyjne metodą taksonomiczną. Praca doktorska. Pol. Śl. Gliwice 1976.
- [7] Pluta W.: O pewnej metodzie klasyfikacji przedsiębiorstw. P.S. 1, 1972.
- [8] Praca zbiorowa pod redakcją Welfe W.: Metody i modele ekonomiczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną. PWE, Warszawa 1981.
- [9] Strahl D.: Propozycje konstrukcji miary syntetycznej. P.S. 2, 1978.
- [10] Siedlecka U.: Zastosowanie metody taksonomii stochastycznej do dyskryminacji zbiorów skończonych. P.S. 4, 1976.
- [11] Więckowski J.: Metody analizy ekonomicznej w przemyśle. PWE, Warszawa 1977.

ИЗБРАННЫЕ ВОПРОСЫ МНОГОМЕРНОГО СРАВНИТЕЛЬНОГО АНАЛИЗА  
ДЛЯ ПРОБЛЕМ ГОРНОГО ДЕЛА

## Р е з ю м е

В статье представлены основные проблемы многомерного сравнительного анализа, обращая при этом особое внимание на приготовление множества переменных описывающих исследуемый объект. Учтён при этом цель анализа, способ нормирования переменных и процедуры выбора. Теоретические рассуждения иллюстрированы практическим примером.

SELECTED PROBLEMS OF MULTI-DIMENSIONAL COMPARATIVE ANALYSIS IN REGARD  
TO MINING PROBLEMS

## S u m m a r y

In the paper are presented the basic problems of multi-dimensional comparative analysis calling attention to the preparation of a set of variables describing the studied object for the purpose of analysis, normalization of variables method, procedure of selecting allocation. The theoretical considerations are supplemented with an example from mining practice.